



U.S. Army Research Institute for the Behavioral & Social Sciences

FACT SHEET



Evaluating Decision Aiding Systems

Purpose

To provide an overview of guidelines for evaluating military decision aiding systems.

Facts

1. The following material is based on the book, *Handbook for Evaluating Knowledge-Based Systems: Conceptual Framework and Compendium of Methods*, by Leonard Adelman and Sharon L. Riedel. It was published in 1997 by Kluwer Academic Publishing. Four-year funding for the book was jointly provided by the U.S. Army Research Institute (ARI) and the U.S. Army Artificial Intelligence Center. The evaluation framework and many of the evaluations methods described in the book were developed and validated using the Airland Battle Management-Advanced Technology Transition Demonstration as a testbed..
2. Automated systems and decision aids have become a part of how the Army does business. The cost of their development in terms of time, manpower, and money is substantial. Life cycle (LC) evaluations which are systematically integrated into the development LC of the system and rigorously implemented will increase the probability that the final system is useful, usable, and used. But this integration will not just happen; the evaluations need to be planned and managed from the outset.
3. One problem with LC evaluations is that standard military procedures for test and evaluation are not entirely appropriate for evaluating systems developed using rapid prototyping. Problems include requirements that are not firm until late

Handbook for Evaluating Knowledge-Based Systems

by Leonard Adelman and Sharon L. Riedel

provides a framework for evaluating decision making systems, identifies the types of evaluations that should be conducted at different

stages of system development, and describes evaluation methods and measures that are appropriate for different types of evaluations at the different system development stages.



in development, unreliable software, the need for rapid evaluation feedback in order to influence a fast moving development process, uncertainties about the capabilities that will be available to evaluate at any point and difficulty in validating the knowledge base when experts do not agree. However, a framework is available which is appropriate for systems developed using rapid prototyping.

4. The ARI Life Cycle evaluation approach identifies what types of evaluations should be conducted at different stages in system development, what issues should be addressed at each stage, and what evaluation methods, measures, criteria and data sources are appropriate for each issue at each stage. Types of evaluations addressed at various points in the development LC include requirements validation, knowledge base validation, knowledge base verification, usability evaluation, and performance evaluation. These might be performed at several stages of development using different methods and different data sources. For example, the



usability of the system could be evaluated at a preliminary design stage by a Human Factors specialist examining planned interface displays. At a later development stage when the system is operational, usability could be evaluated by the number and types of errors made by the intended users when using the system.

For additional information, please contact Dr. Sharon Riedel, ARI - Fort Leavenworth Research Unit, DSN 552-9764 or Commercial (913) 684-9764. riedels@leavenworth.army.mil.